Aalto University
School of Business

The Comparison of Fama-French Five-Factor Model in Chinese A-share Stock Market and in Real Estate Sector

Master’s Thesis
Shao Yufang
Aalto University School of Business
M.Sc. in Finance
Spring 2017

Author Shao Yufang

Title of thesis The Comparison of Fama-French Five-Factor Model in Chinese A-share Stock Market and in Real Estate Sector

Degree Master of Science in Finance

Degree programme Finance

Thesis advisor Professor Vesa Puttonen

Year of approval 2017 **Number of pages** 55 **Language** English

Purpose of the study

This paper aims to test Fama-French five factors model in the Chinese A-share stock markets and in the real estate industry in order to better understand the Chinese markets and to provide the Chinese investors a reliable asset pricing model.

Data and methodology

The data is mainly extracted from Datastream. The tracking period is from July 1st 2002 to December 31st 2015 in total 162 months. There were more than 3000 listed companies in the Chinese A-share market and 138 listed companies in the real estate industry in this study. The methodology is mainly to apply Fama-French five factors model in the Chinese A-share stock markets and in the real estate industry along with other tests such as Durbin-Watson test and multicollinearity to evaluate the model performance.

Findings

The Fama-French five factors model performs well in the Chinese A-share market and the Chinese real estate industry with the data from July 2002 to December 2015. Value factor is not helpful on explaining the excess return of either A-share size-B/M portfolio or real estate industry size-B/M portfolio. The reason why value factor is insignificant in explaining A-share is the speculative environment in Chinese stock market. However, speculation is not the key which makes value factor insignificant in real estate market. Investment factor has limited explanatory power in the A-share size-B/M portfolio too. Surprisingly, five factors model performs better in the Chinese real estate industry than in the Chinese A-share market because five out of six factors are significant at 5% level and the factors are less correlated.

Key words Fama-French, five-factor model, Chinese A-share market, real estate industry

Abstract

Fama-French five factors model performs well in the Chinese A-share market and the Chinese real estate industry from July 2002 to December 2015. The excess return of the A-share size-B/M (Book to Market value) portfolio can be captured by the market excess return, size and operating profitability factors. On the other hand, the excess return of the real estate industry size-B/M portfolio can be captured by the market excess return, size, operating profitability, and investment factors. Value factor is not helpful on explaining the excess return of A-share size-B/M due to the speculative environment in Chinese A-share market. Value factor is not helpful on explaining the excess return of either real estate size-B/M portfolio either. However, the reason is not because of the speculative environment. Investment factor has limited explanatory power in the A-share size-B/M portfolio too. Surprisingly, five factors model performs better in the Chinese real estate industry than in the Chinese A-share market because five out of six factors are significant at 5% level and the factors are less correlated.

Contents

1. Introduction

China is the second largest economy with a GDP of 67,670.8 billion [1]CNY in 2015 right behind USA. Under the Belt and Road Initiative, Chinese corporations invested 14.82 billion USD directly to 49 related countries in 2015. At the same time, Chinese corporations undertook 3,987 contracts, which the total monetary amount reached 92.64 billion USD, with 60 counties.[2] The Asian Infrastructure Investment Bank (AIIB) was founded under the lead of China on 25th December 2015. AIIB is aimed to strengthen the infrastructure in Asia and speed up the economic integration process. Therefore, China has become more and more important in Asia as well as in the world. Thus, it will be very useful to know more about the Chinese stock market and understand how it works. This paper intends to test Fama-French five factors model which works in the western world. I want to know whether this theory works with the Chinese data.

Although China is leading in the emerging market, it is still immature comparing with USA in the matter of stock exchange history. After the founding of the Shanghai Stock Exchange (SHSE) in 1990 and the Shenzhen Stock Exchange (SZSE) in 1991, China started the stock exchange in modern life. Since then, not only the institutional investors have played an important role in the Chinese stock market, but also millions and millions of individual investors have devoted their asset to it. According to Guo Shuqing, the former chairman of China Securities Regulatory Commission, 80% of the trading volume in Chinese stock market were contributed by individual investors in 2013[3]. Here brings the question: what is the percentage of these individual investors having been professionally trained in a trading strategy? Are these individual investors creating a speculative investing environment?

On the other hand, China has experienced a high economic growth rate for more than 30 years. According to the Nation Bureau of Statistic, the GDP growth rate in 2015 was 6.9%[4]. Although the growth rate was lower than the counterpart in 2014, still the growth rate was much higher than the global average which was 2.4%[5]. During the past, it was the manufacturing industry acting as the engine in the Chinese economy. Nowadays, the weight of the tertiary

[1] http://www.stats.gov.cn/tjsj/zxfb/201601/t20160120_1306759.html
[2] http://www.mofcom.gov.cn/article/ae/ah/diaocd/201601/20160101238681.shtml
[3] http://finance.sina.com.cn/stock/y/20130116/185114307116.shtml
[4] http://www.stats.gov.cn/tjsj/zxfb/201601/t20160120_1306759.html
[5] http://www.stats.gov.cn/tjsj/sjjd/201603/t20160309_1328611.html

sector of the economy is the highest in GDP. When the citizens become richer and richer, they start to relocate their asset in other investments such as stock markets or real estates. People buy residential housing for living purpose but also for investment and speculative purposes. Housing has become the top priority of Chinese citizens' life, especially for men. In Beijing, Shanghai, and Shenzhen, it was told that dating men cannot propose to his other part without an apartment in the city. Future mother-in-law will forbid their daughters to marry a man without an apartment under his name, or even more, under the couples' name. Frankly speaking, an apartment in the tier one cities, including Beijing, Shanghai, Shenzhen, and Guangzhou, is representing status and wealth.

From the 2015 February report of China Index Academy, the Shanghai first hand residential housing price index increased 15% in 2015 compared to the number in 2014 and increased 0.25% compared to the number last month[6]. From the same report, the second hand one increased 23% compared to the counterpart last year and increased 1.82% compared to the counterpart last month.[7] According to the news, in February 2016 one second hand apartment seller change his quote for three times in a day with a price rise of 700,000 CNY yet the buyer would like to accept the raise of 500,000 CNY[8]. Although this is an extreme case, it is obvious that the real estate market is facing a fluctuation. Theory that the Chinese real estate bubble will crash soon is widely accepted. In order to forecast the price increase of the real estate in China, as it is such important for Chinese citizens. This paper is going to test the Fama-French five factors model with the real estate market data as well.

To be more precise, I want to figure out the asset pricing model in the stock market in general and in the real estate sector. Eugene Fama and Kenneth French first published a three-factor asset pricing model in 1993. They discovered that for stocks, portfolios constructed to mimic risk factors related to size and BE/ME capture strong common variation in returns, no matter what else is in the time-series regressions (FF, 1993). After years, they developed the three-factor model into a five-factor model in 2015. They drew a conclusion that the five-factor model (FF, 2015) including size, B/M ratios, profitability, and investment, performs better than the three-factor model.

---

[6] http://industry.fang.com/en/default.html
[7] http://industry.fang.com/index/DataIndex.aspx
[8] http://www.yicai.com/news/2016/02/4753110.html

This article aims to test Fama-French five factors model in the Chinese A-share stock markets and in the real estate industry in order to better understand the Chinese markets and to provide the Chinese investors a reliable asset pricing model.

2. Objectives and Innovation

Based on the research of Fama-French, this article is going to test the Fama-French five-factor asset price model's performance in Chinese A-share stock market and in the real estate industry both from SHSE and SZSE. My hypothesis is that Fama-French five-factor model will have statistically significant results on capturing the average return in both Chinese A-share stock market and in the real estate industry. I assume coefficients of the factors in the real estate industry will be similar to the coefficients of the factors in the Chinese A-share stock market.

Hypothesis 1: Fama-French five-factor model will have statistically significant results on capturing the average return in Chinese A-share stock market.

Hypothesis 2: Fama-French five-factor model will have statistically significant results on capturing the average return in Chinese real estate market.

Hypothesis 3: Coefficients of the factors in the real estate industry will be similar to the coefficients of the factors in the Chinese A-share stock market.

There are rather a few innovations in this article. Since the publication of Fama-French three factors model in 1993, many researchers and scholars have tested this model in Chinese markets with different periods and different samples. However, the five factors model of Fama and French just came out in 2015. Little literature has discussed, which I have found, tested the five factors model in the context of the Chinese market.

Another significance of this study is that I try to test the five factors model in a single industry. Both the three factors model and five factors model require to construct a value-weighted portfolio. Fama and French (2015) used all NYSE, AMEX, and NASDAQ stocks to construct this portfolio. Wang and Guo (2014) used 100 largest stocks out of CSI300 index to construct the portfolio yet produced convincing result. Chen, Hu, Shao, and Wang (2015) constructed the value-weighted portfolio with all the non-financial stocks in the SHSE and

SZSE. So far, no one has tried to construct a portfolio in a single industry. The reason might be the firms in the same industries might be affected by the similar outside factors such as policies, cyclical nature, and consumer expectations. The correlation between the firms is strong. This article will test whether the portfolio from the same industries works. The real estate firms were one of the earliest listed firms in SHSE and SZSE. Most of the real estate firms have comparably extensive historical records in our sample period. It makes the result more reliable.

Thirdly, Fama and French (2015) excluded financial firms in their study. Meng and Ju (2013) excluded financial stocks, mega-caps companies in their study. Chen, Demirer, and P. Jategaonkar (2015) excluded utilities firms in their study. They believed those companies will affect the performance of the model. But in this paper, I will include financial firms and utilities firms for the reason that these companies play an important role in Chinese stock market. Excluding them also needs huge adjustments to the index.

Lastly, Wang and Xu (2004) and Chen, Hu, Shao, and Wang (2015) have claimed that the speculative environment is the key that B/M ratio is not functioning well in the Chinese market. But they haven't provide the evidence that how speculative the Chinese stock market is. It is also very difficult to measure speculative level. There isn't any particular speculative metrics from which I can just tell how speculative the investing environment is. In this paper, metrics to measure speculative level as well as to test the speculative level in Chinese A-share market will be presented.

3. Literature review

There are couple questions I would like to solve before testing the five factors model. First of all, how robust is the explanatory power of the three factors or five factors model? Secondly, how does three factors model behave in China? In the end, what is the relationship between the Chinese A-share stock market and the real estate industry?

3.1 Explanatory power of the three factors models

To challenge CAPM and other asset pricing model's explanatory power, Fama and French (1993) addressed that three-factor model, which includes market excess return, size, and B/M factors, explains the average return on stocks and bonds.

They used stocks from NYSE, Amex, and NASDAQ from July 1963 to December 1991. The market excess return has been calculated as the monthly market return minus the risk free rate. The size factor, SMB, has been calculated as the return of the small size group minus the return of the big size group. The value factor, B/M ratio, has been calculated as the return of the high B/M ratio group minus the return of the low B/M ratio group.

The methodology of Fama and French in 1993 is the time-series regression approach developed from Black, Jensen, and Scholes (1972). Monthly returns on stocks and bonds are regressed on the returns to a market portfolio of stocks and mimicking portfolios for size, B/M ratio, and term-structure risk factors in returns. There are other two reasons why they use time-series regression approach. One is that if assets are priced rationally, variables that are related to average returns must proxy for sensitivity to common risk factors in returns. In our case, size and B/M ratio. The other one is that the estimated intercepts in these time-series regressions provides a simple return metric and a formal test of the capability in capturing the cross-section of average return by various combination of the common factors.

The results of Fama and French (1993) include two parts. Firstly, size and B/M ratio are indeed proxy for sensitivity to common risk factors in stock returns because they capture strong common variation in returns, no matter what else is in the time-series regression. Secondly, a market factor, size, and B/M ratio have good explanatory power on the cross-section of average stock returns.

After the publication of three factors model, many other researchers have tested it in different markets. Griffin (2002) suggested that the three factors model performed better on a country-specific basis rather than a global basis. The data Griffin used in this research is that he collected 1521 Japanese companies', 1234 British companies', and 631 Canadian companies' monthly stock return from January 1981 to December 1995. Then he ran three factors model on domestic level, world level, and international level. The adjusted R^2 from the international model is 0.904, which is 0.006 higher than counterpart from the domestic model. However, the pricing error from the international model is 0.24, as compared to 0.22 with the domestic model. On the other hand, the international model yields higher absolute intercepts that means a higher possibility of misspecification. Thus, Griffin drew a conclusion that the three factors model performs better on domestic level than on international level. This research gives us a hint that the three factors model not only functions in the U.S. market but also works in the British, Japanese, and Canadian markets.

3.2 Explanatory power of the three factors models in China

However, these papers are dependent on the data from U.S., U.K., Canada, and Japan. As early as 2004, Chen (2004) had empirical evidence confirming that the common variation of cross-sectional expected stock returns can be explained by size and B/M ratio. What I can learn from this research is that the three factors model functions in the Chinese stock markets too.

Later, Meng and Ju (2013) also found out that the three-factor model had good explanatory power on A-share market of Shanghai Exchange. In their research, they excluded financial stocks, mega-caps companies, and ST stocks (Special Treatment stocks, which indicate the risks are higher than normal stocks.) to prevent high B/M ratio, large-cap, and high volatility from affecting the portfolio too much. But in my research, I keep the financial stocks and mega-caps companies for the reason that financial stocks and mega-caps companies play an important role in the Chinese stock market. Though, excluding these factors? will bring a smoother result, it will bring a result which cannot reflect to whole market. Their sample period rages from July 2005 to June 2012. They only tests with the Shanghai A-shares but not the Shenzhen A-shares. Compared with their data set, I have larger samples and longer tracking period. Meng and Ju (2013) used a traditional time-series regression as Fama and French did. With this method, they have concluded the market factor explains most of the stock return changes, followed by size factor and B/M ratio factor.

In the following year, Xu and Zhang (2014)'s finding also supported this idea. They have concluded that B/P ratio works better in three-factor model rather than BE/ME ratio. The tracking period Xu and Zhang used is from 1991 to 2011. 20 years are quite long as compared with the short history of Chinese stock market. On the other hand, an innovation from them is using B/P ratio instead of B/M ratio for the reason that the Chinese companies are listed on different stock exchanges such as A-shares in mainland, H-shares in Hong Kong, and N-shares in New York. In this case, they thought it would have been incorrect to measure B/M ratio. Furthermore, the Small Medium Enterprise Board (SEB) and Growth Enterprise Board (GEB) have been set up in the Shenzhen Stock Exchange. Thus, they have tested two samples in the Fama-French three factor model. One includes SEB and GEB stocks and one excludes.

As a result, Xu and Zhang have found out the three factors model can explain more than 93% of the variation in the portfolio returns on Chinese A-share market. Secondly, it is better to include SEB and GEB stocks when dividing companies into size group. Thirdly, they recommend to use B/P ratio rather than B/M ratio as a factor.

Liu and Wang (2013) tested the relationship between idiosyncratic risk and expected stock return with Fama-French three factor model. They have drawn conclusions which the size factor is positively correlated with the stock return and the B/M ratio is generally correlated with stock return. They used the daily and monthly data of Chinese A-share market from January 2000 to March 2011. Then Liu and Wang conducted Fama-French three factors model and EGARCH (1,1) model to estimate the idiosyncratic risk. The result has shown that the size, turnover, liquidity, and B/M ratio are positively correlated with the return of stocks. But the momentum factor and the return of stocks are negatively correlated.

Gan, Hu, Liu, and Li (2015) used the data from 1996 to 2005 to contradict the findings above stating that the size factor was negatively correlated with the stock return. They used Chinese A-share market data from January 1996 to December 2005. The risk free rate they used was the fixed deposit rate of the first month of each year. Gan et al. (2015) examined three factors model and CAPM in this research. The result has shown that the high B/M ratio portfolio stock mean return is 0.004 and the low B/M ratio portfolio stock mean return is -0.003. It means that the high B/M ratio portfolio stock is higher than the low B/M ratio portfolio stock mean return. In addition to the mean return, the average standard deviation of the high B/M ratio portfolio is 0.0038 lower than the counterpart of the low B/M ratio portfolio. Based on the findings above, they have confirmed the size and B/M ratio effects exist in the Chinese A-share market. However, the adjusted R^2 of their findings is 0.4195, which is lower than the one from Fama French (1993). In this case, they have drawn a conclusion that the explanatory power of the three factors model is less effective in the Chinese stock market than in the U.S. market.

Lin, Wang, and Cai (2012) conducted their research with the data of SHSE and prove that the Fama-French factors are good proxies for risk factors of portfolio. This is suggesting that Fama-French factors are representative for risk premium in Chinese stock market. The tracking period of Lin, Wang, and Cai's research is from January 2000 to December 2009. They have constructed 100 portfolios but only use 237 individual stock. The method they used was to estimate risk factors from returns of portfolio by principal components. Later on, they compared market factor, size factor, and B/M factor with the estimated risk factors to test the adequacy of using those three factors for risk factors. The result is that all three factors are good proxies for risk factors of portfolio. However, only market factor is good proxy for risk factor of individual stock. My paper has longer tracking period and much larger samples than their research. It is possible that my findings will contradict to theirs.

3.3 History of factors' evolvement

Couple years later, Chen, Demirer, and P. Jategaonkar (2015) overthrew the idea above claiming that equity return dispersion would serve as a more meaningful proxy for risk factor than others in China. They created their own five factors model by adding the return dispersion factor and the idiosyncratic volatility factor into the Fama-French three factors model for capturing the return variation. They used A-share stocks both from SHSE and SZSE for the period July 1996 through June 2011. The proxy of risk free rate they used was three-month household deposit rate. The financial firms were deleted from the sample too. However, Chen et al. did not exclude the mega-caps companies but excluded the utilities and firms that did not have an industry code available.

The methodology in the research is running the cross-sectional regressions to find our which model is the best. As a result, they have found out that the adjusted R^2 of the model with the market factor, size factor, B/M factor, return dispersion factor, and idiosyncratic volatility factor is 0.629, which is the highest as compared to the other three factors or four factors models. In the same five factors model, the portfolios formed on return dispersion and idiosyncratic volatility has an adjusted R^2 as 0.839, which is 0.195 higher than the one of the portfolio formed on size and B/M ratio. Based on their experiments, they have concluded that equity return dispersion captures the fundamental uncertainty that cannot be captured by market and firm level factors.

Drew, Naughton, and Veeraraghavan (2003) suggested if investors in China selectd some combination of small and low book to market equity firms in addition to the market portfolio, they would generate superior risk-adjusted returns. This is a hint indirectly shown to the Chinese investors to follow Fama-French size and B/M ratio factor in order to generate a superior return. Also it indicates that the B/M ratio works oppositely in the Chinese stock market.

The tracking period for the sample is from December 1993 to December 2000. Firstly, the tracking period is shorter compared to many other similar studies. Secondly, it wasn't long after the establishment of Chinese stock exchange. The market wasn't mature at that time. All these may deal effects on the result. Drew et al. use the China 1-Year Time Deposit Rate as the risk free rate of return.

In their research, they have tested not only Fama-French three factors model but also tested the January and Chinese New Year effect by adding seasonal factors. They have created a dummy variable for the January effect and one for the Chinese New Year effect. According to their findings, the small and growth firms generate superior return in China. It means that the value effect isn't as pervasive as was found for the U.S. portfolio and other international markets. This might be very important to our findings.

Hou, Xue, and Zhang (2014) had empirical evidence showing that average stock returns could be summarized by a market factor, a size factor, an investment factor, and a profitability factor. They have developed their own four factors model called q-factor model including the market factor, size factor, investment factor, and profitability factor. Compared with the Fama-French three factors model, the four factors model has excluded the value factor and include two new factors. Hou et al. used stocks from NYSE, Amex, and NASDAQ as sample. The sample period is from January 1972 to December 2012. They calculated the annual change in total assets divided by 1-year-lagged total assets as investment-to-asset. The profitability is measured as the income before extraordinary items divided by 1-quarter-lagged book equity. The methodology they used was fitting the 35 significant anomalies in the broad cross section with different factor models including the four factors model. The result has shown that the q-factor model outperforms in capturing many of the significant anomalies than the Fama-French three factors model and Carhart model do.

Twenty-three years later, Fama and French (2015) developed a five-factor model based on their three-factor model. They have discovered that average stocks return can be better explained by adding two more factors which are operating profitability and investment.

The tracking period is July 1963 through December 2013, which is 264 months longer than the tracking period of their 1993 study. This may affect the comparison of the results in these two studies. The sample is, however, the same from all NYSE, Amex and NASDAQ stocks. The way Fama and French constructed the market factor, size factor, and value factor is the same. On the other hand, they calculated the operating profitability in a way which revenues subtract cost of goods sold, subtract selling, general, and administrative expense, subtract interest expense all divided by book equity. The investment is measured as the change in total assets from the fiscal year ending in year t – 2 to the fiscal year ending in t – 1, divided by t – 2 total assets. It is quite close to the way Hou et al.'s way to measure investment with one year lagged.

Similar to Hou, Xue, and Zhang's finding, Fama and French (2015) also thought HML was a redundant factor. So that they came up with an alternative five-factor model with a factor called HMLO (orthogonal HML) while the rest of factors remaining the same. However, the results of five-factor model with HML factor or HMLO are pretty much the same. In this article, I decide to use the five factors model with HML factor instead of the one with HMLO factor as it is easier to construct.

As for the result, Fama and French haven't given us the adjusted R^2 from the five factors model regression itself. But they give us the coefficients and the t-statistics for these coefficients. According to their saying, the three factors model were affected by both small extreme growth stocks and large extreme growth stocks. Microcap extreme growth stocks alone was a big problem too. The five factors model reduced this problem as the intercept raised 0.2 compared with the counterpart in three factors model. Furthermore, three out of four extreme growth portfolios' intercepts came close to zero. But the small extreme growth stocks still provided negative intercepts and the large extreme growth stocks provided positive intercepts in five factors model.

Based on the former research that the three-factor model works well in the A-share market and that the five-factor model explains better than the three-factor model, I have reason to believe that the five-factor model will work in the A-share market. But there are also research in which the B/M ratios are not robust in China.

Wang and Xu (2004) used all the A-share from July 1996 to June 2002 as sample. They have discovered that the book to market ratio factor is not helpful on explaining the stock returns while the size factor is still functional. The methodology they used was to change the B/M ratio factor into free float factor. Because they tested the B/M ratio factor wasn't statistically significant in all the models but the free float factor is. In their opinions, the free float would affect the companies' future cash flow, which in return a better proxy for Chinese companies' growth potentials and investment opportunities.

According to Wang and Xu's findings, the average adjusted R^2 the time-series regression including market factor, size factor, and free float factor is 0.90. The reason for B/M ratio factor not functioning in China was because most of the Chinese investors were seeking for short term gain, in another word, speculative. Thus, the average trading volume and the participation rate were abnormal in the Chinese stock market. The B/M ratio was useless in capturing the cross-sectional variances in the stock returns.

Chen, Hu, Shao, and Wang (2015) also have discovered that there isn't significant robust value effect on Chinese stock market by using three-factor model with the data from July 1997 to December 2013. According to their conclusions, the reason for this was because the robust value effect was caused by a few extreme months before 1997. I need to consider that the HML factor might not work in our five factors model either.

Chen et al. have extracted all the stocks from both SHSE and SZSE as the research sample. They have employed the Fama-French three factors model regressions and the Fama-Macbeth regressions approaches. The three factors model regressions performed very well in capturing the cross-sectional variations in average returns on portfolios. However, three factors played different roles in this. The most important factor they concluded in the time-series regression was size factor. Unlike many other researches, the B/M ratio factor, on the contrary, was weak in explaining the cross-sectional variations in average returns.

3.4 The relationship between equity market and real estate industry

Most of the researchers have tested the three factors model on whole market bases. Although no one has tried to use Fama French three factors model on a single market, Xie and Qu (2016) used real estate index to represent the market return in their study.

The samples Xie and Qu used were all the A-share stocks from SHSE. The sample period is from January 2005 to December 2012. They were quite innovative on the methodology in this research. The first innovation was the way to construct the factors. Instead of using "total market value" as in Fama-French three factors model, they used "market value in circulation". The reason they did it in this way was that there were many non-tradable in the secondary market such as state-owned corporate shares, state shares, and other non-tradeable shares. These non-tradable shares couldn't be liquidated even though the price shrunk. Furthermore, the companies issued the non-tradable shares were implied a higher degree of nationalization and lower market efficiency. In this case, Xie and Qu had sorted all the A-share stocks according to market value in circulation when they constructed the size factor and size-B/M portfolio. But they insisted using total market value when they calculated the B/M ratio to lower multi-collinearity with other independent variables. The second innovation was that they used three factors model not only to explain the average return of market portfolios but also to explain the average return of sector portfolios. They tested four sector portfolios including industrial sector, commercial sector, real estate sector, and utility sector.

The results of Xie and Qu's research are satisfied. They have got an average adjusted R^2 of 0.8736 in the market portfolio time-series regression. The average adjusted R^2 of industrial sector is 0.9268, followed by the one of utility sector, which is 0.8740. The commercial sector ranks the third with an average adjust R^2 of 0.7884. The last one is real estate factor as 0.6039. The results suggest that the three factors model is applicable in explaining the cross-sectional variation in both market and sector portfolios. Size premium and value premium exist in A-share market and sectors. However, the utility sector is most sensitive to value factor while other sectors are insensitive.

In another study, Zhang and Fung (2006) found that the Chinese stock market and real estate market were negatively correlated. They examined the booming in the housing market could in part explain the bearish of the stock market in the period from 2001 to 2004. This might contradict to the result this paper is expecting to have.

Zhang and Fung gathered the quarterly national house prices and Shanghai house prices from 1997 to the second quarter of 2005 as well as the Shanghai and Shenzhen composite indexes. Then, they ran the multivariate regression to find out the relationship between the stock indexes and the real estate prices. Later on, they tested the Granger causality relationship between the stock prices and the housing prices.

From the multivariate regression results, Zhang and Fung learned that house prices were inversely related to changes in stock prices in SHSE after the listed companies' profitability being controlled. This indicated that housing price had a significant explanatory power to Shanghai stock composite index. But the housing price is an insignificant explanatory variable of Shenzhen stock composite index. This phenomenon suggested that geography led to market segmentation. From the Granger causality Wald test, Zhang and Fung have got a chi-square of 4.97, with p-value of 0.00258. This meant the surge of housing prices had caused the drop of the Shanghai stock composite index. Similar to the multivariate regression results, there wasn't significant Granger causality from the housing prices to the Shenzhen stock composite index. Moreover, the result also shown that there wasn't feedback effect from the changes of stock price to house prices.

On the other hand, Burdekin and Tao (2014) overthrew this saying and claimed that the equity and property market tended to move together in China. They used the data of the Shanghai A-share market index and an index of Shanghai housing prices from November 1999 to May 2011. Beside testing the relationship between share prices and house prices, they linked both series with liquidity, interest rates, loan growth, and overall inflation in order to reflect consumer prices.

The methodology Burdekin and Tao used was the Granger causality Wald test. Contradict to the result of Zhang and Fung (2006), Burdekin and Tao's result revealed significant bidirectional causality between housing price growth and share price growth from 1999 to 2011. The empirical result confirmed that the stock market and housing market move together. In addition to the Granger causality Wald test, Burdekin and Tao also tested a multivariate vector autoregressive framework. In this test, however, there wasn't significant effect of housing prices on share prices. It has confirmed that there is significant effect of share price growth on housing price growth.

Meanwhile, Lin and Lin (2011) analyzed there wasn't a causality relationship between stock and real estate markets in China from the period of March 1995 to June 2010. Their methodology included four parts. First was the unit root test. Second was the cointegration test. Third was the nonlinear cointegration test. Last but not least, was the Granger causality test. The unit root test was to ensure that the series are integrated to the same degree. To examine the market integration, they used Johansen cointegration test. After testing the cointegration between two markets, they tested the fractional integration with nonlinear cointegration test. Finally, Lin and Lin used the Granger causality test to examine the causal relationship between the stock and real estate market.

The result of unit root test has shown that all the sample stock market indexes and real estate indexes are non-stationary. The Johansen cointegration test's result has shown that there isn't cointegration relationship between the Chinese stock and real estate markets. However, the result of nonlinear cointegration test is significantly different from zero and one, which indicating that the Chinese stock market and real estate market are fractional integrated. There isn't any evidence in the Granger causality test rejecting the null hypothesis that the Chinese stock market and real estate market are unrelated.

Zhou and Sornettea (2004) said short-term investors herd for fast gain in Chinese stock market because of the immaturity of the market itself. This finding might imply that the prices of stocks deviate because of short-term investors behavior, which in return affected our result. The data sets Zhou and Sornettea used were the SHSE, SZSE, and TX Investment Co. Ltd. indexes. The tracking period is from August 2000 to October 2003. They used a parametric detrending approach, a generalized q-analysis, and a combined log-periodic power law analysis to find out the 2001 Chinese stock market antibubble. The result has shown that the anti-bubble in the Chinese stock market may be linked to the real estate bubble in China at that period of time. Their findings also suggest that though the Chinese stock market is immature and strongly influenced by the Chinese government, the investors still have herding behavior.

All of the studies discussed above are summarized in the table 1 below.

Table 1

Summary of previous literatures:

Study	Models	Market	Tracking period	Publish time	Notes
Fama and French (1993)	Three factors (market, size, B/M ratio)	U.S.	July 1963 to December 1991	1993	
Griffin (2002)	Three factors (market, size, B/M ratio)	U.K., Japan, Canada	January 1981 to December 1995	2002	1. 1521 Japanese companies, 1234 British companies, 631 Canadian companies 2. R^2 is 0.898
Meng and Ju (2013)	Three factors (market, size, B/M ratio)	China (Only in SHSE)	July 2005 to June 2012	2013	
Xu and Zhang (2014)	Three factors (market, size, B/P ratio)	China	1991 to 2011	2014	1. R^2 is 0.93 2. Replace B/M ratio with B/P ratio
Liu and Wang (2013)	Three factors (market, size, B/M ratio)	China	January 2000 to March 2011	2013	
Gan, Hu, Liu, and Li (2013)	Three factors (market, size, B/M ratio)	China	January 1996 to December 2005	2015	R^2 is 0.42

Lin, Wang, and Cai (2012)	Three factors (market, size, B/M ratio)	China	January 2000 to December 2009	2012	237 companies
Chen, Demirer, and P. Jategaonkar (2015)	Five factors (market, size, B/M ratio, return dispersion, idiosyncratic volatility)	China	July 1996 through June 2011	2015	R^2 is 0.839
Drew, Naughton, and Veeraraghavan (2003)	Four factors (market, size, B/M ratio, season)	China	December 1993 to December 2000	2003	Use China 1-year time deposit rate as risk free rate
Hou, Xue, and Zhang (2014)	Four factors (market, size, investment, profitability)	U.S.	January 1972 to December 2012	2014	B/M ratio factor is redundant
Fama and French (2015)	Five factors (market, size, B/M ratio, operating profitability, investment)	U.S.	July 1963 through December 2013	2015	1. 606 months of tracking period 2. B/M ratio factor is redundant
Wang and Xu (2004)	Three factors (market, size, free float)	China	July 1996 to June 2002	2004	1. B/M ratio factor is insignificant 2. R^2 is 0.9
Chen, Hu, Shao, and Wang (2015)	Three factors (market, size, B/M ratio)	China	July 1997 to December 2013	2015	B/M ratio factor is insignificant
Xie and Qu (2016)	Three factors (market, size, B/M ratio)	China (Only in SHSE)	January 2005 to December 2012.	2016	1. Use real estate index to represent market return 2. R^2 is 0.604

Zhang and Fung (2006)	Multivariate regression and Granger causality test	China	1997 to 2005	2006	Housing price has a significant explanatory power to SHSE composite index
Burdekin and Tao (2014)	Granger causality test	China	1999 to 2011	2014	Stock market and housing market move together
Lin and Lin (2011)	Unit root test, cointegration test, nonlinear cointegration test, Granger causality test	China	March 1995 to June 2010	2011	There isn't causality relationship between stock market and real estate market
Zhou and Sornettea (2004)	Parametric detrending approach, q-analysis, lo-periodic power law analysis	China	August 2000 to October 2003	2004	Investors herd for fast gain in Chinese stock market

Note: The study is referring to the name of the study and its authors. The models mean the methodologies being used in the corresponding study. The market is referring to where the data is extracted. The tracking period means the time period that the study covers. The publish time means when the paper is published. The notes are the important findings of the corresponding study.

4. Data and Experiments

4.1 Data

I extract the data from Datastream. As China A-share stock market has a short history from 1991, I planned to use the data from 1996 so that there would be 240 months data. Unfortunately, Datastream only provides most of the data since January 1st 2000. On the other hand, one of the five factors, investment, requires total asset from year t - 2. This means if year t - 2 can be as early as year 2000. All in all, I will assume year t to start from the year 2002. The tracking period will start from July 1st 2002 to December 31st 2015 in total 162 months. There were more than 3000 listed companies in the Chinese A-share market and 138 listed companies in the real estate industry at the end of 2015. The data includes macro statistics such as one-year Renminbi deposit rate and monthly price on the SHSE index and SZSE index. It also includes micro statistics such as market cap, market to book ratio, revenue, cost of goods sold, selling, general, and administrative expenses, interest expense, book equity, and total asset of each company.

4.2 Factors in A-share market

The Fama-French five factors model is expressed in the equation below:

$$R_{it}-R_{ft}=a_i+b_i(R_{Mt}-R_{Ft})+s_iSMB_t+h_iHML_t+r_iRMWt+c_iCMA_t+e_{it} \quad (1)$$

In equation (1), there are seven variables I need from the database. Rit is the average monthly return of portfolio i. Rf is the risk free rate. Fama and French chose one-month U.S. Treasury bill rate as the risk free rate (FF, 2015). In China, there isn't any bonds issued by the government whose duration is one month until 2007. I need to look for something else for the risk free rate. Meng and Ju (2013) used Chinese central bank bill yield to measure risk free rate but they did not specify the duration of the central bank bill. Wang and Guo (2014) used one-year Treasury bill rate to represent risk free rate which would be higher than the one-month Treasury bill. Learn from their experience, I am going to use one-year deposit rate to represent risk free rate. The reason is one-year Renminbi deposit rate is usually lower than one-year Treasury bill rate so that it can mimic the one-month risk free rate better. Rm is the value-weighted market portfolio return. I am using the weighted average monthly return on the SHSE index and SZSE index to represent the Rm. SMB is the variable related to size. The size group which depends on the market cap of the stocks in A-share market will be separated by median breakpoint of the market cap in A-Share market at the end of June each year. HML is the variable related to book to market (B/M) ratio. The B/M group is depended on the B/M ratio of the stocks in the A-share market. It will also be separated by median breakpoint of the B/M ratio in the A-Share market at the end of June each year. RMW is the variable related to operating profitability. This is not determined by a simple ratio but by a calculation using revenue subtract cost of goods sold, subtract selling, general, and administrative expenses, subtract interest expense and the total divided by book equity. CMA is the variable related to investment. The investment of a company is calculated as the total asset at the end of year t - 1 subtract the one of year t - 2 and all divided by the total asset at the end of year t - 2.

Table 2

Construction of Size, B/M, profitability, and investment factors in 2 X 2 sorts (A-share market): Size group: small(S) or big(B), B/M group: high(H) or low(L), OP group: robust(R) or weak(W), Inv group: conservative(C) or aggressive(A):

Breakpoints	Factors and their components
Size: SHSE and SZSE median	SMB=(SH+SL+SR+SW+SC+SA)/6-(BH+BL+BR+BW+BC+BA)/6 (small minus big)
B/M: SHSE and SZSE median	HML=(SH+BH)/2-(SL+BL)/2 (high minus low B/M)
OP: SHSE and SZSE median	RMW=(SR+BR)/2-(SW+BW)/2 (robust minus weak OP)
Inv: SHSE and SZSE median	CMA=(SC+BC)/2-(SA+BA)/2 (conservative minus aggressive Inv)

Note: All the SHSE and SZSE A-share stocks are allocated in two clusters according to SHSE and SZSE median market cap at the end of each June from 2002 to 2015.

4.3 Factors in real estate industry

In the same principle as allocating the A-share stocks, I will allocate the A-share real estate stocks, which in total 138 stocks, into the portfolio. Risk free rate for the real estate market is the same as in the A-share market which is one-year Renminbi deposit rate. As for Rm, I will use the value-weighted average monthly return on the SHSE real estate index and the SZSE real estate index. The rest of the factors are conducted in the same way as factors of A-share do but with the sample of 138 real estate firms.

Table 3
Construction of Size, B/M, profitability, and investment factors in 2 X 2 sorts (Real estate industry):
Size group: small(S) or big(B), B/M group: high(H) or low(L), OP group: robust(R) or weak(W), Inv group: conservative(C) or aggressive(A)

Breakpoints	Factors and their components
Size: SHSE and SZSE real estate industry median	SMB=(SH+SL+SR+SW+SC+SA)/6-(BH+BL+BR+BW+BC+BA)/6 (small minus big)
B/M: SHSE and SZSE real estate industry median	HML=(SH+BH)/2-(SL+BL)/2 (high minus low B/M)
OP: SHSE and SZSE real estate industry median	RMW=(SR+BR)/2-(SW+BW)/2 (robust minus weak OP)
Inv: SHSE and SZSE real estate industry median	CMA=(SC+BC)/2-(SA+BA)/2 (conservative minus aggressive Inv)

4.4 Experiments

After getting the data, I need to construct the factors. Fama and French (1993) used 2 X 3 sorts for factors. The idea is to intersect two size groups and three B/M groups. Size groups are divided by the NYSE median market cap. B/M groups are divided by the 30th and 70th percentiles of B/M for NYSE stocks. Fama and French (2015) kept the 2 X 3 sorts and further developed the 2 X 2 sorts and the 2 X 2 X 2 X 2 sorts for four factors. Through their tests, average return and standard deviations of SMB are quite close to each version of the factors. Other three factors depend more on the way they are constructed. The 2 X 3 sorts absorb more of the extreme cases while the 2 X 2 sorts tend to be milder. The 2 X 2 X 2 X 2 sorts also have a high correlation with other two versions on HML factor. The difference among three versions of factors construction comes while constructing CMA and RMW factors. The 2 X 2 X 2 X 2 sorts perform better on constructing CMA with 0.37 correlation with HML when the counterpart in 2 X 2 sorts and 2 X 3 sorts is close to 0.70. However, it becomes opposite on constructing RMW factor. The correlation of RMW and HML is 0.04 in 2 X 2 sorts and 0.08 in 2 X 3 sorts when it is 0.63 in 2 X 2 X 2 X 2 sorts. Our goal is to choose one way which reduces the correlation between factors and also produces a significant result to construct the factors. Chen, Hu, Shao, and Wang (2015) found there was significant size effect on Chinese stock market return. This reinforces the value of 2 X 2 sorts as it values higher on the size. In a nut shell, 2 X 2 sorts seem easier to construct yet fits our goal.

When the factors are ready, I need to test these factors are stationary. I will use Augmented Dickey-Fuller test in order to verify whether these factors have unit root or not. The results of the Augmented Dickey-Fuller test are listed in the appendix table 25 and table 26. From the results of Augmented Dickey-Fuller test, I can easily conclude that all the factors from A-share market and real estate industry don’t have unit roots, which means they are stationary. Hence the following regressions shall be meaningful.

Most importantly, I will run the Fama-French five factors time series regression on A-share market and real estate market. After getting the regression result, I need to test Durbin-Watson test to find out whether the error terms are not positively autocorrelated or not. Last but not least, multicollinearity can be a big problem. I am going to test the multicollinearity between the independent variables.

5. Discussions

5.1 Findings in A-share market

In the size-B/M portfolio I construct for A-share market, the average percent return is shown in table 4. The average return of size-B/M portfolios follow a patter. Except for the small group, all the other size groups' percent returns increase against B/M ratio. In another word, the smaller the B/M ratio the higher the return is. Only two portfolios in the small group don't fit this pattern, which means 92% of the size-B/M portfolios in the Chinese A-share market with lower B/M ratio have higher return. This is a quite clear difference from the results found by Fama and French from the U.S. market. On the other hand, only the highest two B/M ratio portfolios follow a pattern that the smaller size brings the higher return. There are eight portfolios don't fit this pattern. Size effect do exist but is not very satisfying in the Chinese A-share market.

Table 4

Average monthly A-share stocks percent returns for portfolios formed on Size and B/M:

	Big	4	3	2	Small
High	0.670	0.840	0.870	0.980	1.240
4	1.060	1.100	1.110	1.130	1.410
3	1.250	1.390	1.320	1.440	1.350
2	1.790	1.920	1.700	1.510	1.490
Low	2.450	2.260	2.470	1.900	1.220

Note: All the SHSE and SZSE A-share stocks are allocated in five clusters according to their market cap at the end of each June from 2002 to 2015. The same stocks are allocated in five clusters according to their book to market ratio at the end of each June from 2002 to 2015. Big to small are referring to five clusters allocated by descending market cap. High to low are referring to five clusters allocated by descending B/M ratio. The intercepts of these five size clusters and five B/M clusters construct a 5 X 5 portfolio matrix. This table shows the average monthly returns of these 25 size-B/M portfolios from July 2002 to December 2015.

The A-share descriptive statistics are listed in Table 5. The means of all five factors' percentage returns are very close to zero. The standard deviations of Rm-Rf, SMB, and RMW's percentage returns are more than twice as big as the counterparts of HML and CMA's percentage returns. This indicates that the variations of Rm-R-, SMB, and RMW's percentage returns are much larger. Furthermore, I can see the operating profitability factor, RMW, can produce up to 0.568% per month.

Table 5

A-share market factors descriptive statistics:

	Mean(%)	Maximum(%)	Minimum(%)	Std. Deviation	N
Rit-Rf	0.397	8.185	-6.001	2.414	162
Rm-Rf	-0.012	0.230	-0.272	0.085	162
SMB	-0.014	0.170	-0.490	0.072	162
HML	0.007	0.096	-0.100	0.028	162
RMW	0.012	0.568	-0.151	0.077	162
CMA	-0.001	0.084	-0.071	0.025	162

Note: Rit-Rf is size-B/M portfolio excess percentage return. Rm-Rf is market excess percentage return. SMB is percentage return of size factor. HML is percentage return of B/M factor. RMW is percentage return of operating profitability factor. CMA is percentage return of investment factor. All the data is from July 2002-December 2015 in total 162 months.

The regression summary is given in Table 6. The results show that the adjusted R square is 0.91, which means the excess return of A-share size-B/M portfolio can be well explained by the market excess return, SMB, HML, RMW, and CMA factors. To test autocorrelation, I used Durbin-Watson test. The lower bound critical value of Durbin-Watson test with sample size of 162 and 5 independent variables at 5% significance level is 1.680 and the upper bound is 1.807. The test result of 1.777 lies between the lower bound and the upper bound, which indicating that the test, the error terms being positively autocorrelated, is inconclusive.

Table 6

The model summary of five factors model on A-share market, July 2002-December 2015, 162 months:

R	R^2	Adjusted R^2	Std. Error of the Estimate	Durbin-Watson
0.956	0.914	0.911	0.720	1.777

Note: The dependent variable is Rit-Rf which is the size-B/M portfolio excess percentage return. The independent variables are α, CMA, Rm-Rf, SMB, HML, and RMW, which α is constant, Rm-Rf is market excess percentage return, SMB is percentage return of size factor. HML is percentage return of B/M factor, RMW is percentage return of operating profitability factor, and CMA is percentage return of investment factor. The adjusted R^2 is 0.911, which indicating 91.1% of the variability of the dependent variable can be explained by the independent variables. The lower bound critical value of Durbin-Watson test with sample size of 162 and 5 independent variables at 5% significance level is 1.680 and the upper bound is 1.807[9]. The test result of 1.777 lies between the lower bound and the upper bound, which indicating that the hypothesis, the error terms being positively autocorrelated, is inconclusive.

However, combining with Table 7, the coefficients statistics for A-share market factors, I can see the details of coefficients and get a more vivid understanding about the regression. Market factor, SMB factor, and RMW factor are significant at 1% level. Both the market factor and operating profitability factor play important roles in explaining the variation of Chinese A-share size-B/M portfolio's return. Because the coefficients of these two factors are the largest. Nevertheless, HML and CMA factors are statistically insignificant. The reason that HML is not significant at 5% level might be coherent with Wang and Xu's finding which suggests the speculative investment environment in China, asymmetrical information, and market friction are the causes (Wang and Xu, 2004). For this reason, I will test the speculative in the discussion of real estate section. It might also be our sample period is after 1997, which is a more stable period so that the value effect is not robust (Chen, Hu, Shao, and Wang, 2015). Or perhaps, HML is just a redundant factor as Fama and French found in the U.S. market (FF, 2015).

[9] http://web.stanford.edu/~clint/bench/dw05b.htm

Table 7

Coefficients statistics for A-share market factors:

Factors	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
α	0.581	0.061		9.542	0.000	0.460	0.701
Rm-Rf	19.755	0.783	0.695	25.231	0.000	18.208	21.301
SMB	14.266	1.658	0.426	8.605	0.000	10.991	17.540
HML	0.603	2.546	0.007	0.237	0.813	-4.427	5.633
RMW	19.667	1.585	0.630	12.405	0.000	16.535	22.798
CMA	-2.190	4.138	-0.023	-0.529	0.597	-10.365	5.984

Note: The dependent variable is Rit-Rf which is the size-B/M portfolio excess percentage return. Rm-Rf is market excess percentage return. SMB is percentage return of size factor. HML is percentage return of B/M factor. RMW is percentage return of operating profitability factor. CMA is percentage return of investment factor. Four out of six independent variables including α, Rm-Rf, SMB, and RMW are statistically significant at 1% level. HML and CMA are statistically insignificant. The standardized coefficients are the changes of the standard deviations of Rit-Rf when the standard deviations of referring independent variables increase 1.

I used Pearson correlation in this research to find out more about the correlations. In terms of the partial correlation as other independent variables being controlled, the market factor and operating profitability factor are the top two predictors correlated with the dependent variable. The partial correlations are 0.896 and 0.705 respectively. The collinearity statistics in table 8 indicates the variance inflation factor and its reciprocal, tolerance. The multicollinearity is not high between the dependent variable and the independent variables though SMB and RMW are relatively higher than other factors. Some scholars argued that if the variance inflation factor is higher than 3 but less than 10, it will suggest these factors have possible collinearity problem[10].

[10] http://statisticalhorizons.com/multicollinearity

Table 8

Correlations between predictors and dependent variable and collinearity statistics for A-share market factors:

Factors	Correlations			Collinearity Statistics	
	Zero-order	Partial	Part	Tolerance	VIF
α					
Rm-Rf	0.868	0.896	0.593	0.728	1.373
SMB	-0.214	0.567	0.202	0.225	4.436
HML	-0.276	0.019	0.006	0.613	1.631
RMW	0.658	0.705	0.292	0.214	4.665
CMA	0.466	-0.042	-0.012	0.298	3.356

Note: The dependent variable is Rit-Rf which is the size-B/M portfolio excess percentage return. The independent variables are α, CMA, Rm-Rf, SMB, HML, and RMW, which α is constant, Rm-Rf is market excess percentage return, SMB is percentage return of size factor, HML is percentage return of B/M factor, RMW is percentage return of operating profitability factor, and CMA is percentage return of investment factor. The correlation statistics shows the Pearson correlation statistics. More importantly, I shall look at the partial correlation as other independent variables being controlled. The collinearity statistics indicates the variance inflation factor and its reciprocal, tolerance. The multicollinearity is not high between the dependent variable and the independent variables, Although, SMB and RMW are relatively higher than other factors.

In order to find out this problem, I conduct a collinearity test and the result is shown in Table 9. The dimensions 1 to 6 are referring to 6 predictors, which are α, Rm-Rf, SMB, HML, RMW, and CMA. Eigenvalues and condition indexes are determining how severe the multicollinearity of the factor is. The closer the eigenvalue to zero or the bigger the condition index is the more severe multicollinearity problem. The lowest eigenvalue and the highest condition index are from CMA factor. It has a eigenvalue of 0.091 and condition index of 4.674. They are still in the safe zone suggesting that all the independent variables do not have severe multicollinearity problem. But I need to keep in mind that CMA has relatively higher multicollinearity than other variables.

Table 9

Collinearity test for A-share market factors:

Dimension	Eigenvalue	Condition Index	Variance Proportions					
			α	Rm-Rf	SMB	HML	RMW	CMA
1	1.995	1.000	0.000	0.090	0.030	0.010	0.040	0.010
2	1.921	1.019	0.050	0.010	0.020	0.100	0.000	0.050
3	0.999	1.413	0.610	0.060	0.010	0.000	0.000	0.050
4	0.541	1.920	0.170	0.820	0.040	0.070	0.040	0.010
5	0.453	2.099	0.170	0.000	0.040	0.690	0.030	0.140
6	0.091	4.674	0.010	0.010	0.870	0.130	0.880	0.740

Note: The dimensions 1 to 6 are referring to 6 predictors, which are α, Rm-Rf, SMB, HML, RMW, and CMA. α is constant. Rm-Rf is market excess percentage return. SMB is percentage return of size factor. HML is percentage return of B/M factor, RMW is percentage return of operating profitability factor. CMA is percentage return of investment factor. Eigenvalues and condition indexes are determining how severe the multicollinearity of the factor is. The closer the eigenvalue to zero or the bigger the condition index is the more severe multicollinearity problem. Normally if the eigenvalue less than 0.01 or the condition index larger than 20, there is a severe multicollinearity problem. In this case, all the independent variables do not have severe multicollinearity problem. CMA, has relatively higher multicollinearity than other variables.

In table 10, I can see that SMB has a correlation of 0.837 with RMW and -0.712 with CMA. It means that SMB is highly positively correlated with RMW and highly negatively correlated with CMA. As I discussed above, there was a high correlation between RMW and CMA factors in the 2 X 2 sorts. SMB factor is highly correlated with RMW and CMA. This makes the three factors are highly correlated. Unfortunately, I haven’t tried the 2 X 2 X 2 X 2 sorts in this article. Perhaps in that case, these factors will have a lower correlation. It would be a good idea for future study.

Table 10

Coefficient correlations for A-share market factors:

		CMA	Rm-Rf	SMB	HML	RMW
Correlations	CMA	1.000	-0.016	-0.712	0.513	-0.697
	Rm-Rf	-0.016	1.000	0.013	0.187	-0.239
	SMB	-0.712	0.013	1.000	-0.163	0.837
	HML	0.513	0.187	-0.163	1.000	-0.242
	RMW	-0.697	-0.239	0.837	-0.242	1.000
Covariances	CMA	17.125	-0.051	-4.885	5.407	-4.573
	Rm-Rf	-0.051	0.613	0.017	0.372	-0.296
	SMB	-4.885	0.017	2.748	-0.688	2.201
	HML	5.407	0.372	-0.688	6.483	-0.977
	RMW	-4.573	-0.296	2.201	-0.977	2.513

Note: CMA is percentage return of investment factor. Rm-Rf is market excess percentage return. SMB is percentage return of size factor. HML is percentage return of B/M factor, RMW is percentage return of operating profitability factor. This table shows the correlation and covariance between these variables. SMB and RMW are the most positively correlated. Two pairs of variables including both CMA-SMB and CMA-RMW are the most negatively correlated.

5.2 Findings in real estate industry

In the size-B/M portfolio I construct for real estate industry, the average return is shown in table 11. Unlike the A-share size-B/M portfolio, the value effect has changed in real estate industry. It seems like the higher the B/M ratio, the higher the portfolio return will be. There are five portfolios do not fit in this pattern. On the other hand, size effect in real estate market is a little bit better than the one in A-share market. The smaller size portfolio tends to produce a higher return. This pattern applies to 19 out of 25 portfolios. From the productivity perspective, the small-high real estate portfolio produce the highest average return of 4.22% per month, which is 1.75% higher than the 3-low portfolio, the highest portfolio in A-share market.

Table 11

Average monthly real estate stocks percent returns for portfolios formed on Size and B/M; July 2002-December 2015, 162 months:

	Big	4	3	2	Small
High	1.63	1.47	1.91	2.01	4.22
4	1.42	1.72	1.79	2.00	1.92
3	1.06	0.58	0.83	1.57	4.61
2	1.24	0.28	1.38	1.83	1.26
Low	-0.17	-1.18	0.49	1.50	1.79

Note: All the SHSE and SZSE A-share real estate stocks are allocated in five clusters according to their market cap at the end of each June from 2002 to 2015. The same stocks are allocated in five clusters according to their book to market ratio at the end of each June from 2002 to 2015. Big to small are referring to five clusters allocated by descending market cap. High to low are referring to five clusters allocated by descending B/M ratio. The intercepts of these five size clusters and five B/M clusters construct a 5 X 5 portfolio matrix. This table shows the average monthly returns of these 25 size-B/M portfolios from July 2002 to December 2015.

The real estate industry descriptive statistics are shown in Table 12. The means of the variables much different as compared with zero. Especially, SMB factor has a mean return of

0.594% per month. The standard deviations are much larger than the ones in A-share markets. This is suggesting that the percentage returns of the factors in real estate market vary a lot.

Table 12

Real estate factors descriptive statistics:

Factors	Mean(%)	Maximum(%)	Minimum(%)	Std. Deviation	N
Rit(RE)-Rf	0.406	8.126	-6.142	2.682	162
Rm(RE)-Rf	-0.007	0.365	-0.282	0.114	162
SMB	0.594	15.308	-9.686	3.071	162
HML	0.340	10.684	-8.304	2.598	162
RMW	-0.056	6.529	-7.561	2.269	162
CMA	-0.274	6.134	-20.519	2.763	162

Note: Rit(RE)-Rf is size-B/M real estate portfolio excess percentage return. Rm(RE)-Rf is real estate market excess percentage return. SMB is percentage return of real estate size factor. HML is percentage return of B/M real estate factor. RMW is percentage return of real estate operating profitability factor. CMA is percentage return of real estate investment factor. All the data is from July 2002-December 2015 in total 162 months.

The regression summary is listed in Table 13. Surprisingly, the adjusted R^2 is 0.857. Regardless of the fact that the adjusted R^2 of the real estate industry is a bit lower than the number of A-share, it is still impressive that the excess return of real estate size-B/M portfolio can be well explained by the market excess return, SMB, HML, RMW, and CMA factors.

Table 13

The model summary of five factors model on real estate industry, July 2002-December 2015, 162 months:

R	R^2	Adjusted R^2	Std. Error of the Estimate	Durbin-Watson
0.928	0.862	0.857	1.013	2.037

Note: The dependent variable is Rit(RE)-Rf which is the size-B/M real estate portfolio excess percentage return. The independent variables are α, CMA, Rm(RE)-Rf, SMB, HML, and RMW, which α is constant, Rm(RE)-Rf is real estate market excess percentage return, SMB is percentage return of size factor. HML is percentage return of B/M factor, RMW is percentage return of operating profitability factor, and CMA is percentage return of investment factor. The adjusted R^2 is 0.857, which indicating 85.7% of the variability of the dependent variable can be explained by the independent variables. The lower bound critical value of Durbin-Watson test with sample size of 162 and 5 independent variables at 5% significance level is 1.680 and the upper bound is 1.807[11]. The test result of 2.037 is larger than the upper bound, which indicating that there isn't statistical evidence that the error terms are positively autocorrelated.

Same as in A-share market, I have conducted Durbin-Watson test for testing autocorrelation. The lower bound critical value of Durbin-Watson test with sample size of 162 and 5 independent variables at 5% significance level is 1.680 and the upper bound is 1.807. The test result of 2.037 is larger than the upper bound, which indicating that there isn't statistical evidence that the error terms are positively autocorrelated.

To combine with table 14, only HML factor is not significant at 5% level. The explanatory power of other factors is significant at 5% level. The real estate market factor plays an important role in explain the variation of real estate portfolio return. Unlike in the A-share market, the real estate operating profitability factor has a negative coefficient of -0.128. It is indicating that the real estate operating profitability factor move against the portfolio return. Again, I can see HML factor is not significant at 5% level. The reason might be same as our finding in A-share market, which is speculative investors. I have conducted a test to find out how speculative the Chinese investors are. The test is in the next section. To compare with the coefficients with the A-share factors, the changes of real estate size-B/M portfolio rely heavily on the changes of market return and move oppositely against operating profitability factor. On the other hand, the changes of A-share size-B/M portfolio are explained by the market factor, size factor, and operating profitability factor on similar weight. The changes of A-share size-B/M portfolio move simultaneously with all the significant factors.

[11] http://web.stanford.edu/~clint/bench/dw05b.htm

Table 14

Coefficients statistics for real estate industry factors:

Factors	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
α	0.435	0.084		5.145	0.000	0.268	0.602
Rm(RE)-Rf	24.024	0.837	1.018	28.696	0.000	22.370	25.678
SMB	0.290	0.037	0.332	7.760	0.000	0.216	0.364
HML	0.022	0.048	0.021	0.452	0.652	-0.073	0.117
RMW	-0.128	0.049	-0.108	-2.588	0.011	-0.226	-0.030
CMA	0.142	0.039	0.147	3.681	0.000	0.066	0.219

Note: The dependent variable is Rit(RE)-Rf which is the size-B/M real estate portfolio excess percentage return. Rm(RE)-Rf is real estate industry excess percentage return. SMB is percentage return of size factor. HML is percentage return of B/M factor. RMW is percentage return of operating profitability factor. CMA is percentage return of investment factor. Five out of six independent variables including α, Rm(RE)-Rf, SMB, RMW, and CMA are statistically significant at 5% level. Only HML statistically insignificant. The standardized coefficients are the changes of the standard deviations of Rit(RE)-Rf when the standard deviations of referring independent variables increase 1.

I can see the variance inflation factors in table 15. All the factors' variance inflation factors are all lower than 3. It means the collinearity is not so likely exist among these factors. But I will run the same collinearity test for real estate industry factors.

Table 15

Correlations between predictors and dependent variable and collinearity statistics for real estate industry factors:

Factors	Correlations			Collinearity Statistics	
	Zero-order	Partial	Part	Tolerance	VIF
α					
Rm(RE)-Rf	0.829	0.917	0.854	0.705	1.419
SMB	0.061	0.528	0.231	0.484	2.065
HML	0.222	0.036	0.013	0.407	2.456
RMW	-0.127	-0.203	-0.077	0.507	1.974
CMA	-0.141	0.283	0.110	0.559	1.788

Note: The dependent variable is Rit(RE)-Rf which is the size-B/M portfolio excess percentage return. The independent variables are α, CMA, Rm(RE)-Rf, SMB, HML, and RMW, which α is constant, Rm(RE)-Rf is real estate market excess percentage return, SMB is percentage return of size factor, HML is percentage return of B/M factor, RMW is percentage return of operating profitability factor, and CMA is percentage return of investment factor. The correlation statistics shows the Pearson correlation statistics. More importantly, I shall look at the partial correlation as other independent variables being controlled. The collinearity statistics indicates the variance inflation factor and its reciprocal, tolerance. The multicollinearity is not high between the dependent variable and the independent variables.

Same as in the A-share market, if the eigenvalue is less than 0.01 or the condition index is larger than 20, it will be considered as severe multicollinearity. From table 16, I can see all the independent variables do not have multicollinearity problem.

Table 16

Collinearity test for real estate industry factors:

Dimension	Eigenvalue	Condition Index	Variance Proportions					
			Constant	Rm(RE)-Rf	SMB	HML	RMW	CMA
1	2.731	1.000	0.00	0.04	0.03	0.04	0.04	0.04
2	1.200	1.508	0.43	0.00	0.07	0.02	0.02	0.07
3	0.862	1.780	0.22	0.29	0.05	0.00	0.16	0.04
4	0.629	2.083	0.24	0.42	0.05	0.00	0.01	0.32
5	0.323	2.906	0.00	0.25	0.22	0.42	0.58	0.04
6	0.254	3.277	0.10	0.00	0.58	.52	0.20	0.49

Note: The dimensions 1 to 6 are referring to 6 predictors, which are α, Rm(RE)-Rf, SMB, HML, RMW, and CMA. α is constant. Rm(RE)-Rf is real estate market excess percentage return. SMB is percentage return of size factor. HML is percentage return of B/M factor, RMW is percentage return of operating profitability factor. CMA is percentage return of investment factor. Eigenvalues and condition indexes are determining how severe the multicollinearity of the factor is. The closer the eigenvalue to zero or the bigger the condition index is the more severe multicollinearity problem. Generally, if the eigenvalue less than 0.01 or the condition index larger than 20, it indicates severe multicollinearity. In this case, all the independent variables do not have multicollinearity problem.

There is a surprising discovery in Table 17 that the factors are much less correlated in the real estate industry than the ones in the A-share market. The highest correlation is between SMB and RMW, which is 0.558. In the A-share market, the highest correlation is 0.837 between SMB and RMW too. From the correlation perspective, Fama-French five factors model performs better in the real estate industry than in the A-share market. Because the coefficient correlations between factors are lower. The factors are more independent, which in return more reliable.

Table 17

Coefficient correlations for real estate industry factors:

		CMA	SMB	Rm(RE)-Rf	RMW	HML
Correlations	CMA	1.000	0.237	0.128	0.173	0.523
	SMB	0.237	1.000	0.198	0.558	0.266
	Rm(RE)-Rf	0.128	0.198	1.000	0.142	-0.290
	RMW	0.173	0.558	0.142	1.000	-0.158
	HML	0.523	0.266	-0.290	-0.158	1.000
Covariances	CMA	0.001	0.000	0.004	0.000	0.001
	SMB	0.000	0.001	0.006	0.001	0.000
	Rm(RE)-Rf	0.004	0.006	0.701	0.006	-0.012
	RMW	0.000	0.001	0.006	0.002	0.000
	HML	0.001	0.000	-0.012	0.000	0.002

Note: CMA is percentage return of investment factor. SMB is percentage return of size factor. Rm(RE)-Rf is real estate market excess percentage return. HML is percentage return of B/M factor, RMW is percentage return of operating profitability factor. This table shows the correlation and covariance between these variables. SMB and RMW is the most positively correlated factors with correlation of 0.558.

Overall, Fama-French five factors models performs better in the real estate industry in China than the A-share market from July 2002 to December 2015.

5.3 Speculative measurement

Wang and Xu (2004) claim speculative environment is the reason why B/M ratio doesn't work in Chinese market. However, they haven't proved the speculative level in the Chinese market. I gather the investors data from China Securities Depository and Clearing Corporation Limited in table 18. There are 99,105.4 thousand registered investors in the market. 98,821.5 thousand investors are individual investors. It means that 99.71% of the investors are individual investors. As we mentioned in the beginning, 80% of the trading volume is contributed by them in 2013. Furthermore, there are 26,112 thousand new investors coming into the market in 2015. The market tends to believe the institutional investors have more information than the individual investors so that the institutional investors will make more rational choices than the individual investors do. From my perspective, the number of individual investors and the trading volume contributed by the individual investors are indications that the Chinese stock market is quite speculative. But I would like to have a precise metric which can evaluate the perspective level.

Table 18

Equity investors data in 2015 (in thousand)

Year	Total investors	New investors	Individual investors	New individual investors	Institutional investors	New institutional investors
2015	99,105.4	26,161.8	98,821.5	26,112.0	283.8	49.6

Note: Total investor means the register investors in China Securities Depository and Clearing Corporation Limited. New investors mean the number of register investors in 2015 minus the number of register investors in 2014. Individual investor means that the investor is a natural person. New individual investors mean the number of individual investors in 2015 minus the number of individual investors in 2014. Institutional investor means that the investor could be a securities firm, a fund or an insurance company etc.. New institutional investors mean the number of Institutional investors in 2015 minus the number of Institutional investors in 2014.

However, I went through many readings there isn't any specific metrics or ratios, which can measure the speculative level. The closest speculative measure would be the one made in Mei, Scheinkman, and Xiong (2004). According to their definition, the speculative measure is determined by the volatility of the difference in beliefs among the Chinese investors about the firm's fundamental value and by the float of the A-shares, among other variables. Though they have this definition of the speculative measure, they haven't calculated in their study. They assume the speculative component is move together with the A-share turnover rate and the speculative component in B-share is 0.

In this case, I need to come up with my own speculative measure. Firstly, I need to define what is speculative. Speculative is characteristic that the investors are focus on short term gains or arbitrage opportunities. Benjamin Graham (1949) writes in his book to define what investment and speculative are. Investment is one which produces continuously adequate return upon through analysis. The operations not meeting these requirements are speculative. Speculative investors might invest based on gossips, rumors, "inside information", or even the voice in their heads. They believe only they know this information which can lead them outperform the others. However, others believe in the same thing. In this case, investors are herding to invest. The most direct link to reflect herding shall be stock trading volume. I gather the trading volume data for a stock or index all year long and expect around half of the trading volumes are above the average and the other half are below the average. If the days that its trading volume is above the average are less or much less than half of the time, I call these days “the herding days”. The herding days are representing the investors are herding to trade, which in return contribute a large amount of trading volumes. I will define the herding days as speculative.

In table 19, I collected the data of SHSE Composite Index and SZSE Composite Index’s daily trading volume. Both indexes follow the ascending order each year. In the same fiscal year, there is a large gap between the maximum and minimum trading volume. This gap can be as large as 28 times of the minimum trading volume.

Table 19

Descriptive statistics of SHSE Composite Index and SZSE Composite Index's daily trading volume:

	Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
SHSE Composite Index (Million shares)	Max	85	47	45	49	104	208	176	276	263	169	180	244	772	857
	Min	3	5	6	7	14	31	30	65	43	40	43	64	56	147
	Mean	10	13	16	16	43	101	66	138	107	87	78	111	173	416
SZSE Composite Index (Million shares)	Max	317	1944	2401	2968	5661	8976	8511	13342	13878	9729	12184	15037	34451	52013
	Min	150	217	310	402	821	1397	1325	3372	3116	2556	3134	5082	5354	10811
	Mean	489	563	873	1051	2358	4856	3127	7029	6598	5171	5662	8984	12459	28238

Note: This table shows the maximum, minimum, and average daily trading volume of SHSE Composite Index and SZSE Composite Index from 2002 to 2015. Though the absolute numbers of trading volume are very different from SHSE Composite Index and SZSE Composite Index, both of them follow the ascending order each year. There is also a large gap between the maximum and minimum trading volume. This gap can be as large as 28 times of the minimum trading volume.

There are the speculative measures in table 20. It shows the days which its trading volume is above the average trading volume of that fiscal year in SHSE and SZSE. In these 14 years' data, the skewness of speculative measures in SHSE is 0.358, the counterpart in SZSE is 0.267. Both of the skewness are positive suggesting that there are much more days crowding below the average trading volume. A small amount of days contributes a large amount of trading volumes. In the extreme case, half of the trading volume in both stock exchanges were contributed by just one third of the trading days. In addition to this, I can see there is a pattern in the speculative days. Investors are herding in a period of time. In some of the months, the trading volume is above the average in every single day and in some of the months, each day's trading volume is lower than the average.

Table 20

Speculative Measures in SHSE and SZSE:

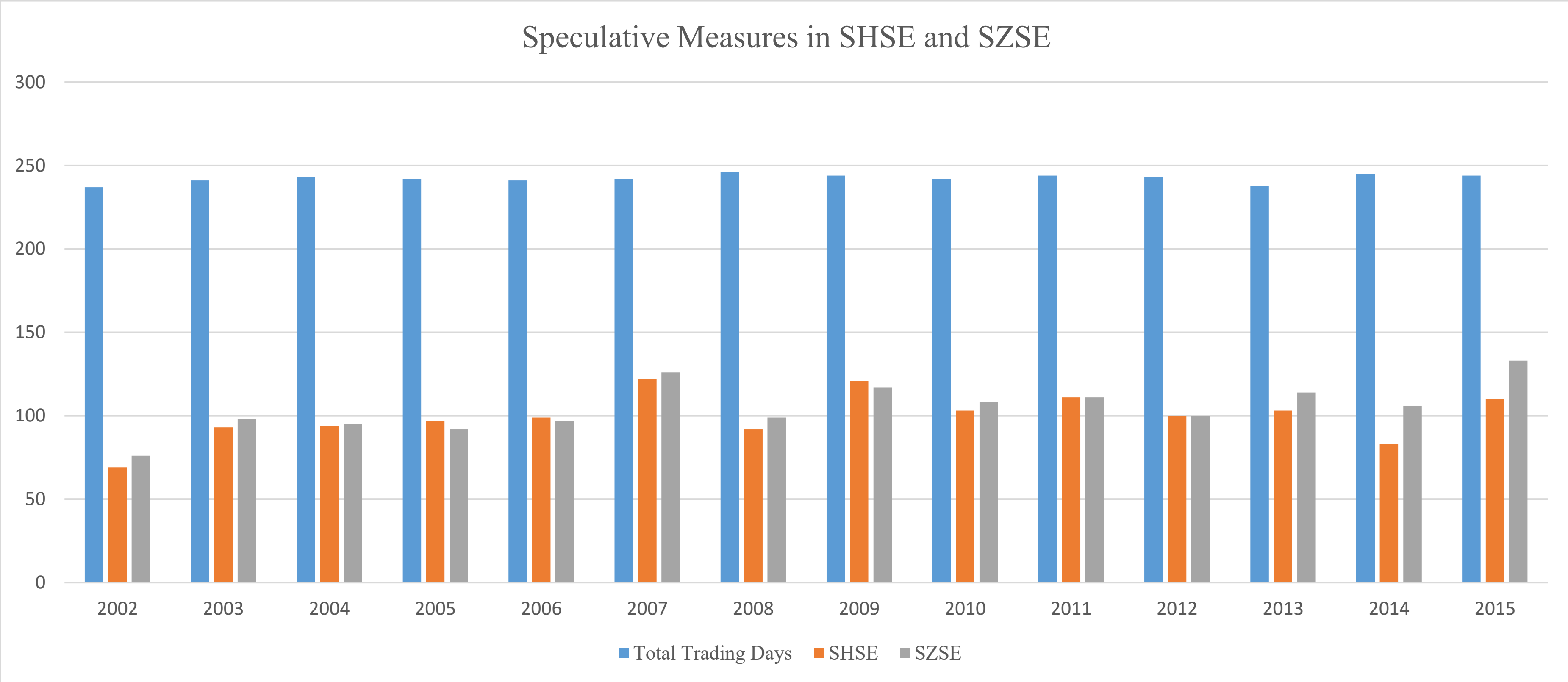
Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
SHSE (days)	69	93	94	97	99	122	92	121	103	111	100	103	83	110
SZSE (days)	76	98	95	92	97	126	99	117	108	111	100	114	106	133
Total Trading days	237	241	243	242	241	242	246	244	242	244	243	238	245	244

Note: This table shows the number of days in which SHSE Composite Index and SZSE Composite Index's trading volume is above the average trading volume of that fiscal year. For example, the average trading volume of SHSE Composite Index is 10 million shares per day in 2002. There are 69 days in which the trading volume is above 10 million shares in 2002.

In the histogram 1, it shows that the speculative measures in both stock exchanges tend to move together. SZSE seems to have relatively more days above the average than SHSE does. It means that investors are more speculative in SHSE than in SZSE. On the other hand, the speculative days have never exceeded 100 from 2002 to 2006. But they are getting closer and closer to the medium. Even more, there are 133 out of 244 trading days are above the average in SZSE. This is the first time that more days are contributing half of the trading volumes in a year. Based on the test I conducted above, I can conclude that the Chinese investors were quite speculative from 2002 to 2006. The overall trend is getting less speculative. Because of the speculative environment, the B/M ratio is not functioning correctly in Chinese A-share market and the real estate industry.

Histogram 1

Speculative measures in SHSE and SZSE:



Note: This histogram show the speculative measures in SHSE and SZSE. The speculative measures in both stock exchanges tend to move together. SZSE seems to have relatively more days above the average than SHSE does. It means that investors are more speculative in SHSE than in SZSE.

Because the overall trend of speculative level in the Chinese market is getting less and less speculative. In order to confirm my findings, I divide the tracking period into two sub-periods based on the speculative measures. The first period is from July 2002 to June 2009. The second period is from July 2009 to December 2015. The reason why I divide the tracking period like this is most of the speculative measures are above 100 days after 2009. I run the Fama-French five factors regression with the A-share data in these two sub-periods. The regression result is listed below in Table 21 and Table 22. The first period is which I consider to be more speculative than the later one. As the table shows, the B/M ratio factor is insignificant at 5% level with the A-share market data from July 2002 to June 2009. However, the B/M ratio factor is significant at 5% level with the A-share market data from July 2009 to December 2015. This result confirms my assumption that the speculative environment is the reason why B/M ratio is not functioning properly in the A-share market.

Table 21

Coefficients statistics for A-share factors in sub-period from July 2002 to June 2009:

Factors	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
α	0.501	0.077		6.533	0.000	0.349	0.654
Rm-Rf	14.264	1.515	0.515	9.417	0.000	11.255	17.273
SMB	-0.500	4.357	-0.007	-0.115	0.909	-9.157	8.156
HML	-6.062	3.221	-0.067	-1.882	0.063	-12.461	0.338
RMW	35.148	4.781	0.577	7.351	0.000	25.649	44.646
CMA	-6.232	5.807	-0.069	-1.073	0.286	-17.768	5.305

Note: The dependent variable is Rit-Rf which is the size-B/M portfolio excess percentage return. Rm-Rf is market excess percentage return. SMB is percentage return of size factor. HML is percentage return of B/M factor. RMW is percentage return of operating profitability factor. CMA is percentage return of investment factor. Three out of six independent variables including α, Rm-Rf, and RMW are statistically significant at 5% level. SMB, HML, and CMA are statistically insignificant at 5% level. The standardized coefficients are the changes of the standard deviations of Rit-Rf when the standard deviations of referring independent variables increase 1.

Table 22

Coefficients statistics for A-share factors in sub-period from July 2009 to December 2015:

Factors	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
α	0.451	0.081		5.553	0.000	0.289	0.614
Rm-Rf	21.570	1.242	0.717	17.369	0.000	19.086	24.054
SMB	9.674	2.093	0.466	4.622	0.000	5.488	13.861
HML	18.900	3.373	0.242	5.603	0.000	12.153	25.648
RMW	13.722	2.178	0.687	6.302	0.000	9.366	18.077
CMA	12.607	4.928	0.112	2.558	0.013	2.750	22.464

Note: Note: The dependent variable is Rit-Rf which is the size-B/M portfolio excess percentage return. Rm-Rf is market excess percentage return. SMB is percentage return of size factor. HML is percentage return of B/M factor. RMW is percentage return of operating profitability factor. CMA is percentage return of investment factor. All six independent variables including α, Rm-Rf, SMB, HML, RMW, and CMA are statistically significant at 5% level. The standardized coefficients are the changes of the standard deviations of Rit-Rf when the standard deviations of referring independent variables increase 1.

I run the Fama-French five factors regression with the real estate market data in these two sub-periods. The regression result is listed below in Table 23 and Table 24. Surprisingly, HML are significant at 5% level in both sub-periods. The reason for this may be that I used SHSE and SZSE composite indexes to represent the whole market speculative level. However, in reality, the investors may not speculatively invest in the real estate stocks. In this case, my hypothesis that the speculative level causes the B/M ratio malfunction in five factors model is rejected. The reason why B/M ratio doesn’t work in five factors model with Chinese real estate data is yet to be discovered. This is one of my research limitation.

Table 23

Coefficients statistics for real estate market factors in sub-period from July 2002 to June 2009:

Factors	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
α	0.290	0.099		2.934	0.004	0.094	0.486
Rm(RE)-Rf	23.089	0.956	0.986	24.143	0.000	21.189	24.989
SMB	0.276	0.045	0.282	6.062	0.000	0.185	0.366
HML	0.211	0.061	0.179	3.447	0.001	0.090	0.333
RMW	-0.216	0.059	-0.163	-3.656	0.000	-0.333	-0.099
CMA	0.253	0.046	0.277	5.538	0.000	0.162	0.344

Note: The dependent variable is Rit(RE)-Rf which is the size-B/M real estate portfolio excess percentage return. Rm(RE)-Rf is real estate industry excess percentage return. SMB is percentage return of size factor. HML is percentage return of B/M factor. RMW is percentage return of operating profitability factor. CMA is percentage return of investment factor. All six independent variables including α, Rm(RE)-Rf, SMB, HML, RMW, and CMA are statistically significant at 5% level. The standardized coefficients are the changes of the standard deviations of Rit(RE)-Rf when the standard deviations of referring independent variables increase 1.

Table 24

Coefficients statistics for real estate market factors in sub-period from July 2009 to December 2015:

Factors	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
α	0.586	0.131		4.456	0.000	0.323	0.849
Rm(RE)-Rf	23.771	1.602	0.992	14.837	0.000	20.566	26.976
SMB	0.245	0.066	0.343	3.688	0.000	0.112	0.377
HML	-0.281	0.091	-0.339	-3.081	0.003	-0.464	-0.099
RMW	0.146	0.097	0.153	1.508	0.137	-0.048	0.340
CMA	0.039	0.079	0.034	0.498	0.620	-0.119	0.197

Note: The dependent variable is Rit(RE)-Rf which is the size-B/M real estate portfolio excess percentage return. Rm(RE)-Rf is real estate industry excess percentage return. SMB is percentage return of size factor. HML is percentage return of B/M factor. RMW is percentage return of operating profitability factor. CMA is percentage return of investment factor. Four out of six independent variables including α, Rm(RE)-Rf, SMB, and HML are statistically significant at 5% level. RMW and CMA are statistically insignificant at 5% level. The standardized coefficients are the changes of the standard deviations of Rit(RE)-Rf when the standard deviations of referring independent variables increase 1.

5.4 Limitations

There are certain limitations in this research. I used one-year Renminbi deposit rate to mimic the risk free rate. It will be a little higher than the one-month Treasury bill rate, which in return leads to inaccuracy. Furthermore, compared with the sample base of Fama and French (2015), which contains all the stocks from NYSE, AMEX, and NASDAQ in total 606 months, our sample base only used about 3000 A-share stocks and 138 real estate stock in total 162 months. Both from the size and from the time length, our data base is small. This might lead to the deviation from fact. From the methodology perspective, I only use 2 X 2 sorts factors in this paper. This might be the reason why A-share market factors are highly correlated. Moreover, I develop the speculative measure based on Mei, Scheinkman, and Xiong (2004)'s usage of turnover. This method might be not ideal. I got the confirmation that the B/M ratio doesn't work in the A-share market because of speculative environment. But I haven't figure out why B/M ratio doesn't work in the real estate market either. This could be a potential topic for future studies.

6. Conclusions

Based on my findings, the Fama-French five factors model performs well in the Chinese A-share market and the Chinese real estate industry with the data from July 2002 to December 2015. The excess return of the A-share size-B/M portfolio can be captured by the market excess return, size, and operating profitability factors. On the other hand, the excess return of the real estate industry size-B/M portfolio can be captured by the market excess return, size, operating profitability, and investment factors. Value factor is not helpful on explaining the excess return of either A-share size-B/M portfolio or real estate industry size-B/M portfolio. The reason why value factor is insignificant in explaining A-share is the speculative environment in Chinese stock market. However, speculation is not the key which makes value factor insignificant in real estate market. Researchers can try to find out why value factor is insignificant in the future. Investment factor has limited explanatory power in the A-share size-B/M portfolio too. Surprisingly, five factors model performs better in the Chinese real estate industry than in the Chinese A-share market because five out of six factors are significant at 5% level and the factors are less correlated.

7. Reference

Burdekin, R. C., & Tao, R. (2014). Chinese real estate market performance: Stock market linkages, liquidity pressures, and inflationary effects. Chinese Economy, 47(2), 5-26.

Chen, C. D., Demirer, R., & Jategaonkar, S. P. (2015). Risk and return in the Chinese stock market: Does equity return dispersion proxy risk?. Pacific-Basin Finance Journal, 33, 23-37.

Chen, C., Hu, X., Shao, Y., & Wang, J. (2015). Fama-French in China: Size and Value Factors in Chinese Stock Returns. University of Hong Kong Working Paper.

Chen Zhan-hui (2004). Cross-Sectional Variations and Three Factors Asset Pricing Model: Empirical Evidences from China A Share Market [J]. Chinese Journal of Management Science, 6, 003.

Drew, M. E., Naughton, T., & Veeraraghavan, M. (2003). Firm size, book-to-market equity and security returns: Evidence from the Shanghai Stock Exchange. Australian Journal of Management, 28(2), 119-139.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. the Journal of Finance, 47(2), 427-465.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of financial economics, 33(1), 3-56.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. Journal of financial economics, 116, 1–22.

Gan, C., Hu, B., Liu, Y., & Li, Z. (2013). An empirical cross-section analysis of stock returns on the Chinese A-share stock market.

Griffin, J. M. (2002). Are the Fama and French factors global or country specific?. Review of Financial Studies, 15(3), 783-803.

Guo, J., & Wang, K. (2014). Empirical tests of Fama-French three-factor model and Principle Component Analysis on the Chinese stock market.

Hou, K., Xue, C., & Zhang, L. (2014). Digesting anomalies: An investment approach. Review of Financial Studies, hhu068.

Lin, T. C., & Lin, Z. H. (2011). Are stock and real estate markets integrated? An empirical study of six Asian economies. Pacific-Basin Finance Journal, 19(5), 571-585.

Lin, J., Wang, M., & Cai, L. (2012). Are the Fama–French factors good proxies for latent risk factors? Evidence from the data of SHSE in China. Economics Letters, 116(2), 265-268.

Liu, Y., & Ping, W. (2013, July). Model selection and relationship between idiosyncratic volatility and expected stock returns: evidence from Chinese A-share Market. In 2013 10th International Conference on Service Systems and Service Management (pp. 522-526). IEEE.

Mei, J., Scheinkman, J. A., & Xiong, W. (2004). Speculative trading and share prices: An analysis of Chinese AB share premia. Working paper (New York University and Princeton University).

Meng, Z., & Ju, R. (2013, December). Explanatory power of three-factor model on A-share market of Shanghai Exchange in China. In 2013 International Conference on Advances in Social Science, Humanities, and Management (ASSHM-13). Atlantis Press.

Wang, C. (2004). Relative strength strategies in China's stock market: 1994–2000. Pacific-Basin Finance Journal, 12(2), 159-177.

Wang, F., & Xu, Y. (2004). What determines Chinese stock returns?. Financial Analysts Journal, 60(6), 65-77.

Xie, S., & Qu, Q. (2016). The Three-Factor Model and Size and Value Premiums in China’s Stock Market. Emerging Markets Finance and Trade, 52(5), 1092-1105.

Zhang, G., & Fung, H. G. (2006). On the imbalance between the real estate market and the stock markets in China. Chinese economy, 39(2), 26-39.

Zhang, S., & Xu, J. (2013). The fama-french three factors in the Chinese stock market.

Zhou, W. X., & Sornette, D. (2004). Antibubble and prediction of China's stock market and real-estate. Physica A: Statistical Mechanics and its Applications, 337(1), 243-268.

8. Appendix

Tabel 25

Augmented Dickey-Fuller test statistics in A-share factors:

Factors	T value	Unit root	Stationary
Rit-Rf	-10.635	No unit root	Yes
Rm-Rf	-10.278	No unit root	Yes
SMB	-8.486	No unit root	Yes
HML	-12.045	No unit root	Yes
RMW	-9.123	No unit root	Yes
CMA	-12.592	No unit root	Yes

Note: Augmented Dickey-Fuller test is to test whether the factors have a unit root or not. The null hypothesis is that the factor has a unit root. When the sample size is 162, the tabulated critical value at 95% confidence level is -3.438. If the T value is lower than -3.438, it means that I can reject the null hypothesis at 5% level, in another word, no unit root. In addition, I will accept the alternative hypothesis that the time series of this factor is stationary.

Table 26

Augmented Dickey-Fuller test in real estate factors:

Factors	T value	Unit root	Stationary
Rit(RE)-Rf	-10.190	No unit root	Yes
Rm(RE)-Rf	-10.509	No unit root	Yes
SMB	-11.790	No unit root	Yes
HML	-14.217	No unit root	Yes
RMW	-13.388	No unit root	Yes

CMA	-13.814	No unit root	Yes

Note: Augmented Dickey-Fuller test is to test whether the factors have a unit root or not. The null hypothesis is that the factor has a unit root. When the sample size is 162, the tabulated critical value at 95% confidence level is -3.438. If the T value is lower than -3.438, it means that I can reject the null hypothesis at 5% level, in another word, no unit root. In addition, I will accept the alternative hypothesis that the time series of this factor is stationary.